P*h*RMA/FDA/AASLD Drug-Induced Hepatotoxicity White Paper
Postmarketing Considerations

November 2000

Note to Readers: This White Paper is divided into two major sections. The first section articulates the current 'state of the art' of postmarketing surveillance for drug-induced hepatotoxicity. The second section addresses the charge of the steering committee to the working group to identify topics and issues worthy of further attention and discussion at the workshop to be held on February 12-13, 2001.

This document was prepared by an FDA Working Group in consultation with representatives of PhRMA and AASLD. It is not a guidance document. It does not contain recommendations to sponsors or applicants regarding particular actions they should take. In stead it is a concept paper that assesses the current state of knowledge, and the existing methodology for examining hepatotoxic events associated with pharmaceuticals. It is meant to provide a framework for discussion at a public workshop on drug-induced hepatotoxicity to be held February 12-13, 2001. As the knowledge evolves about this topic, the Agency may decide to develop guidance on how sponsors can better identify drugs that cause hepatotoxicity during the pre-clinical, clinical, and post-marketing periods.

SECTION 1
Background and Scope of the Issue: Limitations of Pre-clinical and Clinical Safety Databases

Acute hepatic injury is one of the most common forms of pre-clinical and clinical toxicity encountered with drugs. Fortunately, most drugs with the potential for substantial hepatotoxicity will be identified during early testing. Yet, pre-approval testing cannot be relied upon to identify less common drug-related hepatotoxicity because of the recognized and accepted limitations of the safety database submitted in support of a drug licensing application. Phase 3 clinical trials are designed, powered, and reviewed to detect relatively common dose-related adverse drug reactions (ADR) whether or not these are predicted from the drug's pharmacology. However, a database of several thousand patients, which is typical of most marketing applications, lacks the power to observe rare adverse reactions (i.e. those occurring at rate of 1 per 1000 exposures or less [1]. Furthermore, short trial duration and, in some cases, specific patient selection criteria contribute to the inability of pre-marketing studies to detect the rare, drug-related adverse events that may occur when the drug is given to a larger numbers and a broader range of patients following approval. ADRs with a long induction period will not be detected in brief trials and adverse reactions resulting from specific interactions with personal characteristics, concurrent diseases and concomitant drug therapies will not be found.

The size of the pre-approval safety database is not necessarily uniform for all drugs. Although the International Conference on Harmonization (ICH) has issued general guidance on the size and breadth of the requisite safety data, prior experiences with drugs in the same class may warrant special attention or suggest more specific safety investigations. For example, for a chronically administered drug intended for a non-life threatening situation, a clinical safety database of 1500 total subjects of which 300 have been treated for at least 6 months and 100 for one year is expected [1]. However, these numbers may be expanded significantly if there has been unusual toxicity with a pharmacologically or chemically similar drug, or if there are suggestions of potential toxicity in the population studied. The pre-marketing exposure may need to be larger if acceptable safe and effective alternative therapies are already on the market and available. Breakthrough drugs that address unmet medical needs or offer considerable significant improvement over current therapy could be approved more rapidly and with a smaller safety package.

As pre-approval studies are neither designed nor are they able to detect rare, serious events such as severe hepatotoxicity, it is critically important that new information appearing postmarketing regarding drug product safety be collected, analyzed, and acted upon in an expeditious manner. One of the main objectives of post-marketing safety surveillance is to detect rare adverse events associated with the use of drugs, to assess the likelihood of a causal relationship of the event to the drug and, when a drug cause seems likely, estimate the rate of the event. This responsibility is shared by the regulatory authorities, drug manufacturers and by physicians, whose timely reports of adverse drug events are the main source of information about rare adverse effects. It should also be

noted that, in the recent experiences of most serious hepatotoxicity observed postmarketing with specific drugs (e.g. troglitazone, bromphenac), there have been suggestions of less severe liver effects (e.g. elevated transaminases, etc.) in the pre-marketing clinical trials. Additional research should be focused on these specific examples in an attempt to understand the sensitivity/specificity and predictive values of clinical laboratory evaluations. Also, there must be particular attention paid to such suggestions (e.g. transaminase elevation alone or in combination with other clinical laboratories) in drugs being evaluated for approval (see Clinical White Paper).

Current Mechanisms/tools for Identifying and Quantifying Hepatotoxicity of Marketed Drugs

Voluntary spontaneous reporting systems are, undoubtedly, the most cost- effective (despite their limitations) approach for post-market identification of new adverse drug reactions (ADRs). The goal of the evaluation of 'spontaneously' reported information is to provide a signal of new problems or of an increased incidence of problems. Surveillance is conducted on the entire population of treated patients and is not limited by *a priori* hypotheses. Unfortunately, there are major limitations to the spontaneous reporting system the major one of which is under-reporting. Only a variable proportion of adverse drug events cases that have occurred in temporal association with a drug are reported by health care professionals. The proportion of cases reported is always unknown and may vary considerably by for reasons it is impossible to know, making comparisons with other agents or even with known background rates of events difficult. Such factors include severity and novelty of the event, perceived relationship of the event to the drug, adverse effects of related drugs, physician awareness, time since the launch of the drug, previous reports of an adverse reaction, and media interest that has heightened awareness in patients and physicians. In general, these problems do not interfere with its function as a signaling system, but they make rate estimates very difficult.

Reporting rates calculated on the basis of spontaneously reported data can be misleading and are generally presumed to underestimate the risks associated with drug treatments. On the other hand, spontaneously reported events include those whose temporal association with the product drug may be coincidental and differential reporting can make two drugs look very different when they are not. With the exception of some drug-specific diseases or symptoms such as fixed eruption, the risk in unexposed patients (background risk) is never zero, so that reports of a drug association may be incorrect, and reflecting only background occurrence of the event.

The main drawback of spontaneous reporting is the absence of an adequate control group of unexposed patients and, consequently, the difficulty of knowing the relative risk of treated patients, compared to untreated, (unless the event occurs extremely rarely in the population).

Confirmation and Quantification in Denominator-based Systems

Several epidemiological approaches to post-marketing surveillance and quantification techniques have been developed within the United States, some of which, in theory at least have control groups and might allow relative risk estimates.

-Population-based claims databases:
These are large databases that can link prescription-dispensing information (e.g., drug, dose, duration, and date) to claims data for procedures or other medical interventions via unique patient identifiers. These systems can give event rates (numerators) for the denominator of exposed persons, person-years of exposure, etc. and allow comparison with unexposed patients. The limitations of the systems are their ability to record the events of interest and the reliability of the reported events. The availability of medical records in some of these systems, however, allows for validation of event reports as well as more in depth analysis of adverse events and may make formal nested case-control studies feasible [2]. For very rare events, these databases, while large, may still be too small to have adequate numbers of exposures to drugs of interest.

-United Kingdom Prescription Event Monitoring (PEM) Program
In this system, all prescriptions for newly marketed drugs in the United Kingdom are retrieved from the British Prescription Pricing Authority and entered into an automated database. Prescription data supplied to the Drug Safety Research Unit (DSRU) include the name of the patient and the prescribing general practitioner (GP). Subsequently, questionnaires ('green cards') are sent to each prescribing GP 6 months following the first prescription for each patient. Questionnaires request adverse events during and after treatment and reasons for discontinuation, patient age, indication for treatment, starting and stopping dates of treatment, and dosage. Questionnaires are customized for each drug and additional questions can be added [2].

-Exposure Registries
Another cohort technique is the exposure-specific registry, which attempts to identify and follow patients who have been exposed to a drug but have not yet had an adverse outcome (e.g., a pregnancy registry created before delivery). These registries may also be constructed in disease populations (e.g. cystic fibrosis patient registry). The major advantages of such systems are a denominator of exposures. Sometimes registries provide background rates for unexposed (or differently exposed) patients. Registries also may allow more complete follow-up of incident cases.

-Other Systems
A recently developed resource is the Acute Liver Failure (ALF) Study Group, which was begun in January 1998 as a registry and serum bank for all forms of acute liver failure. Data are currently being collected at 14 sites around the U.S. at time of hospital admission and at discharge on patients fulfilling acute liver failure criteria (coagulopathy and altered mentation). To date, the registry has enrolled 150 patients. Preliminary analysis suggests that 50% are related to drug hepatotoxicity (acetaminophen toxicity in 32%) and

idiosyncratic reactions in 18%. Detailed information is available on clinical features, other drugs involved, etc. An analysis of 44 acetaminophen cases is underway. The group is currently expanding to 20 sites and will include pediatric patients. The virtues of this program are detailed data specifically including drug exposure data, and serum samples on severe hepatotoxicity from representative transplant centers around the U.S. The group provides detailed historical information as well as clinical and laboratory features and outcome, including transplantation [W. Lee; Personal Communication, October 1999].

Limitations of These Methodologies

These aforementioned resources have limitations common to most population databases. For example, the PEM has a practitioner response rate to survey of approximately 60%. Rates are provided for reported events only and there is no control group. Limitations of administrative claims systems include the imprecise nature of data linkage of drugs to events; mapping of verbatim event terms to adverse event dictionaries can be problematic, and the temporal relationships between drugs and events are not well established. In most claims data, unless an event leads to hospitalization, the event will not be captured. Some may capture office visit events and sometimes use of a drug is taken as an indicator of an event, obviously a very indirect method. Also, claims interpreted as representing adverse drug events may not actually represent the actual clinical event or outcome of interest (e.g. not all claims of 'seizure' in such databases are actually episodes of epilepsy and are often found to be events of syncope, vasovagal episodes, etc.). All these limitations prevent precise quantification of the incidence of drug related adverse events. The limitation of the ALF Study Group effort is that it does not yet capture all cases of serious hepatotoxicity. No community hospital cases are identified unless they are transferred to a major transplant facility and the group does not currently gather data on cases that are severe but do not reach criteria for fulminant or acute liver failure. These will not be important limitations, however, if it can be shown that the ALF cases are typical of cases in other locations, or of patients experiencing the spectrum of drug-induced hepatotoxicity.

Although the different tools for risk identification and risk quantification have limitations, they probably provide a better picture together than any single method alone. Care should be taken to use and interpret data from these different sources in the proper perspective. The use of several methods and data sources yielding consistent results provides strong support for the hypothesis that the hepatic injury is, indeed, drug related.

Risk/Benefit Assessment of Drug-Related Hepatotoxicity

Having made the best possible assessment of the likelihood that a drug causes severe liver, it becomes necessary to weigh this risk against the drug's value. Apart from drugs treating lethal diseases, any significant rate of severe, potentially fatal liver injury, will lead to a substantial limitation of a drug's use, if not drug withdrawal from the market. This has been true for rates of severe injury at the 1/50,000 or greater level.

It should be appreciated, however, that abnormal liver chemistries, sometimes severe, can arise from many causes (acute and chronic viral infections, alcohol consumption, over the counter and herbal medication use, comorbid disease, and no apparent cause). Virtually all pharmaceutical agents will therefore be associated with severe liver function abnormalities and be reported. Many of these episodes may be unrelated to the drug but it can be difficult to distinguish from true drug-induced cases. Once clinical review of suspected cases (i.e. case series review) is undertaken and, where possible, epidemiological data are reviewed, the list of true hepatotoxic drugs will be narrowed significantly. Once a drug is proven to be seriously hepatotoxic and its toxicity characterized, the critical next step is a careful evaluation of its risks and benefits, and its proper use .

As indicated, drugs with serious hepatotoxicity at a fairly high rate, i.e. 1/10,000 – 1/50,000, like drugs causing aplastic anemia or fatal skin disease at similar rates, will generally be unsuitable for relatively well-treated conditions with good, safer, alternative therapy. For serious illness with limited options, especially where some people do not respond to the alternatives, the risk may be accepted. In order to establish a foundation for decision making, the hepatotoxic risk should be quantified. The most rapid, although assumption-laden method is to compare a calculated reporting rate to a background rate of the hepatic injury of interest. The reporting rate is derived from the number of received reports meeting a case definition over a denominator of drug usage data. This approach illustrates the critical importance of accurate drug utilization data that provides information broken down by demographics (e.g. age, sex, etc.) and setting (inpatient, outpatient).

Another approach is to determine the *excess* incidence of the event associated with drug using data from active comparator clinical trial data, cohort observational studies, case-control studies, or administrative database studies. One or more of these methodological approaches may be used to enhance the robustness of the risk estimate.

Weighing benefits and risks is complex, because benefits and risks are usually expressed in different units [3]. How much risk of fatal liver injury is acceptable for a drug that is an NSAID?, a second line treatment of Parkinson's Disease?, an effective anticonvulsant?, an oral hypoglycemic whose real clinical benefit is often delayed by years to decades? Where decisions are close ones, outside input (FDA Advisory Committees) is often sought.

When severe hepatotoxicity is found for a drug with satisfactory alternatives, the involved drug will usually be removed from the market. Historically, these withdrawals have been voluntary on the part of the manufacturer/sponsor although there are regulatory mechanisms for removal of marketed drugs (Imminent Hazard, Notice of Opportunity for a Hearing on proposed withdrawal).

Where benefits are felt to outweigh the risk of hepatotoxicity, at least in some cases, the more common regulatory risk management tool involves amending the product label to include information on the frequency, severity, clinical indicators, patient prognostic

variables and other information regarding drug-induced hepatotoxicity. The placement of language for label revisions involving hepatotoxic potential has, in the past, been considered on a case-by-case basis; however, recent actions signal a more consistent regulatory approach to prominence and content of this information, usually consisting of bolded or boxed WARNINGS followed by a more detailed discussion of the risk in the WARNING or PRECAUTIONS sections. In almost all cases, there will also be a Patient Package Insert (PPI) or even a "Medication Guide" to allow the patient to participate both in the decision to use the drug and in any needed laboratory or clinical monitoring.

Whether recommendations for liver monitoring in product labeling are a useful and effective risk management tool deserves further discussion. Most often, such recommendations are made based on an intuitive belief that the strategy will allow detection of subclinical liver damage before the process becomes irreversible. This belief is largely based on findings from clinical trials, which show that patients in clinical trials, who are monitored regularly and demonstrate an elevation of aminotransferases above a specified level, will often have a return to normal of their tests after discontinuation of the drug. This observation, however, may not hold for the small subset of patients who progress to fulminant hepatic failure (acute liver failure). In fact, it is clear that in some cases patients may progress from normal aminotransferase levels to fulminant hepatic failure within 1-2 weeks, the usual monitoring interval, after as few as two doses, despite discontinuation of the offending drug. These patients are unlikely to benefit from any conventional or practical liver monitoring strategy. On the other hand, the value of routine and regular liver monitoring is generally accepted as having value in detecting subclinical and reversible liver injury before it progresses to more severe, clinically apparent hepatitis for drugs such as INH.

Much more research is needed to assess the role of liver monitoring strategies in preventing drug induced hepatotoxicity and, if proven valuable, the optimal liver monitoring schedule (see suggested initiatives below). Compliance with liver monitoring recommendations has, not unexpectedly, been shown to be low (May 10 1999 FDA Endocrinologic and Metabolic Drugs Advisory Committee, www.fda.gov/ohrms/dockets/gc/99/transcpt/3499t). It seems clear that monitoring will not be effective unless the patient is fully informed and a partner in the process. The availability of PPIs/Medication Guides and other information provided to the patient are tools that may improve compliance. Finally, even if liver function monitoring is effective in averting drug induced liver failure, the cost effectiveness of the procedure must be considered. Thoughtful research is needed to assess these issues (the value of liver monitoring, the optimal liver monitoring schedule, methods to improve compliance with liver monitoring, the pharmacoeconomics of liver monitoring, etc.).

Significant labeling changes should be accompanied by a mailing ("Dear Dr./Dear Healthcare Professional correspondences") from the pharmaceutical company to the relevant cohort of prescribers and dispensers of this drug (21 CFR 200.5). The effectiveness of such correspondences in modifying physician prescribing and practice behavior has long been a subject of debate. All such information is also made available on

FDA Internet websites and through FDA and industry press releases to news organizations. More research is needed to quantify the impact of these informational risk management strategies.

An intermediate step between product withdrawal and relabeling is the restriction of access to drug. This strategy has been employed in a successful program established to limit access to an effective neuroleptic that may cause a blood dyscrasia to only those patients who undergo regular hematological testing (i.e. clozapine registry). There has also been a recent risk management strategy of limiting the initial use of an effective, parenteral antibiotic (trovofloxacin) to inpatient situations and only in severe infections. Post-approval recognition of the potential of this drug to cause hepatic failure significantly altered its risk-benefit profile necessitating restriction to severe infections.

Another risk-management tool that has been employed is the post-approval informed consent. The hope has been that this strategy allows for a more formalized benefit-risk discussion between prescriber and patient and of the responsibilities of each to insure the safe and effective use of the drug. Again, research is required to quantify adherence to this strategy and its effectiveness in managing risk.

Summary

The previous sections have outlined the current methods of detecting and evaluating drug-induced hepatotoxicity in the post-approval period. The remaining sections consider future directions and recommendations for further research that are offered as topics for discussion at the FDA/P*h*RMA/AASLD workshop.

SECTION 2
Future Directions and Recommendations for Discussion at the Workshop:

Initiative 1: Development of a Standardized Nomenclature and Definitions for Liver Test Abnormalities and Liver Injury as well as a Standardized Approach to Causality Assessment in Postmarketing Reports

Recent experience with evaluation of case series of drug-associated hepatic injury reveals a lack of agreement on approaches to defining and classifying cases as well as widely divergent processes in evaluating these cases for determining causality. There is a critical need for consensus in these areas. The following is an attempt to define the spectrum of drug-induced disease, the parameters and patterns of clinical laboratories that define the injury, and an approach to assessing causality. It is not intended to be definitive or comprehensive but represents a point of departure for further discussion at the workshop.

Any liver test greater than the population-defined upper limit of normal is "abnormal". Liver injury has been defined as any increase to more than 2-3N in ALT or conjugated bilirubin (CB) greater than 2N. Hepatitis is a term reserved for the histologic description of liver pathology (biopsy or autopsy), described as infiltration of mononuclear cells that may or may not be associated with hepatocellular changes [4,5]. The types of liver injuries and their characteristic clinical laboratory patterns are shown in Table 1. Liver injury has been described as acute when liver test abnormalities last three months or less and chronic when these last longer.

Severity is measured by the degree of the laboratory abnormalities, e.g. larger abnormalities of transaminase (10 and > times ULN), and may be accompanied abnormalities of excretory function (hyperbilirubinemia or jaundice, an ominous occurrence) and by effects on synthetic function (prolongation of prothrombin time and reduction of Factor V to less than 50%). Fulminant liver injury is a term used to describe the rapid development, within days to weeks, of hepatic encephalopathy and severe coagulation disorders. This condition frequently leads to cerebral edema, sepsis and death.

Viral etiology is the most frequently identifiable cause of acute failure injury with drugs as the second most common cause [6,7]. The overall incidence of drug related liver injury in most series of acute liver is estimated at 10-20%; it accounts for about 50% of cases among those over the age of 50 years.

The relationship of liver injury to drugs is a critical determination. Causality is highly probable when there is unintentional positive rechallenge phenomenon. Intentional rechallenge is generally considered unethical, as it may be life threatening for the patient. When done, rechallenge is positive if the ALT (hepatocellular injury) or AP (cholestatic injury) is doubled regardless of date, duration, or combination with other uninterrupted drugs; it is negative if the elevation of ALT is less than the upper limit of normal (N)

provided that the drug is given in the same dose, for the same duration, and with the same combined drugs as with the first episode. Liver injury due to drug is probable if other causes could be ruled out, and possible or unlikely under other circumstances. If information is available after the cessation of drug (e.g. dechallenge), drug-induced hepatocellular injury is suggested if the decrease of ALT is more than 50% of the excess over the upper limit of normal within 8 days and no additional elevation of ALT within a month. It is suggestive if the decrease is more than 50% within 30 days, and not suggestive if the variation in ALT levels are otherwise [5,6].

Causal relationship with a drug can only be excluded when the timing is incompatible - onset of liver damage is before the drug is administered, or when the liver abnormalities are discovered months to years after the drug is withdrawn. Underlying illnesses may mimic drug-induced hepatotoxicity and must be considered in attempting to deduce causality. These include but are not limited to chronic alcoholism (elevated aminotransferase with AST to ALT ratio of more than 2 is suggestive of alcohol liver damage), bacterial infection (elevated alkaline phosphatase or total bilirubin with rare elevations of aminotransferase levels above 5N), right sided or biventricular heart failure, and left sided heart failure or hypotension. Right-sided heart failure may lead to congestion of the liver with subsequent very large elevations of AP and/or transaminases, and/or unconjugated bilirubin in most cases. Left-sided heart failure or hypotension, such as occurs as a result of arrhythmia or myocardial infarction, may result in hypoxia of the liver. In such cases, a rapid rise in aminotransferase levels followed by a rapid return to normal is typical of some cases and sometimes accompanied by a delayed hyperbilirubinemia by 48-72 hours. Viral etiologies must be ruled out since this is the most common cause for hepatocellular injury. Fatty liver of pregnancy (cholestatic injury) and biliary obstruction should likewise be eliminated as a possible etiology by hepatobiliary ultrasound.

Action Item: A consensus should be reached on standardized definitions of the entities that make up the spectrum of drug-induced liver disease. These definitions should include the clinical characteristics, clinical laboratory patterns, diagnostic tests, and time courses that differentiate the entities. There should also be an attempt to reach consensus on the approach to causality assessment in reported cases of drug-induced hepatotoxicity.

Initiative 2: Passive surveillance: improvements, increased and improved reporting, and data analysis techniques

The Adverse Event Reporting System (AERS) of the FDA is a passive surveillance system that collects information regarding drug related adverse events that are initially reported by health care providers through either the pharmaceutical companies (about 90% of reports) or directly to the FDA MedWatch program. This voluntary reporting system is largely a paper-based system that requires data entry personnel to enter reports. A transition is underway to convert to direct electronic reporting from pharmaceutical companies to streamline processing of reports. Since its inception in 1993, the MedWatch program has encouraged reporting by increasing health care provider awareness of the MedWatch program through its extensive MedWatch partners network and through publications and presentations in different scientific and health care forums. These efforts have, to some degree, been successful in improving the quality and quantity of reports of serious adverse drug reactions. These efforts should continue and every effort should be made to educate medical students, physicians-in-training and other healthcare professionals on the importance of reporting such events. Education regarding what information should be included in such reports should be emphasized. To augment reporting of liver injury, a checklist for MedWatch is developed to assist the reporter and the MedWatch operator in reporting the critical clinical and laboratory data elements necessary for causality and disease outcome evaluation (see Table 2) [Personal Communication; C Kwong, 1997]. Although this type of checklist may benefit from further development, it is offered as a concept that could be incorporated into the standard operating procedures (SOPs) for pharmaceutical company safety units.

In addition to improving the quality and quantity of reports, the FDA has organized an internal working group to explore the utility of more sophisticated statistical data mining techniques to improve the ability to detect signals of serious adverse drug reactions, including hepatotoxicity, essentially by flagging rates of events that are more than expected, earlier in the post-marketing lifecycle. Preliminary retrospective analyses performed in the former passive surveillance system, the Spontaneous Reporting System (SRS), that collected data from 1969 through October of 1997, suggest that such signaling systems may identify possible adverse effects for hands on safety evaluation by post-marketing staff [8]. Transferring the techniques and knowledge gained from SRS to AERS, an Oracle-based relational system, will require significant time and resource allocation over several years to achieve the promise of an automatic drug ADR signal generation system.

Action Item: Efforts to improve the quality and quantity (yield) of reports on drug-induced hepatic injury should be discussed at the workshop. In addition, a new method of evaluating spontaneous reports is offered as a discussion topic.

Initiative 3: Improved Ascertainment of Background Rates of Elevated Aminotransferases and Hepatic Injury

Interpretation of spontaneous reports of drug-induced hepatotoxicity requires estimates of the background rate of the hepatotoxicity of interest. In contrast to the extensive literature on the clinical, pathologic, and metabolic aspects of drug-induced liver disease, relatively little has been published on its epidemiology [9]. Acute liver failure has been estimated to occur at a rate of approximately 2000 cases per year in the U.S. [7], but little information has been published to document the incidence of clinically important liver disease associated with many commonly used drugs. The few population-based studies that have been published, however, suggest that the incidence of serious drug-induced liver disease among outpatients taking commonly used drugs is low [7,10]. Furthermore, there is scant information regarding the prevalence and types of liver disease in patients with a variety of disorders, such as diabetics. There is a need for a better and more complete understanding of disease- or condition-specific background rates in order to provide a useful context when evaluating drug-associated adverse reactions. Some of the information available includes the following.

Beard et al identified 12 hospitalizations for liver disorders judged possibly or probably attributable to use of outpatient medications other than anti-cancer drugs among 280,000 members of a managed care organization during a five-year period from 1977 through 1981[11]. This corresponds to an incidence of approximately one per 10^5 person-years (py) of exposure within the organization. Walker and Cavanaugh found three cases of new-onset drug-induced hepatitis and 17 new-onset cases of liver disease of uncertain cause during the year 1989 among 71,000 adult members of a managed care organization, yielding incidences of $4/10^5$py and $24/10^5$py, respectively [12]. None of the three drug-related cases was hospitalized and, in the cases of uncertain cause, there were no concurrent prescription medications or blood products that appeared related to the liver injury. In a study of non-steroidal anti-inflammatory drug (NSAID) users in Saskatchewan, Canada, Garcia Rodriguez et al found that hospitalizations for liver disease occurred at an incidence of $10/10^5$py of use, which would correspond to less than 1 hospitalization per 10^5 thirty-day prescriptions[13]. This study made use of a database that permits linkage of all prescriptions and all hospitalizations for residents of the province of Saskatchewan, with a population of approximately one million. All these studies are based on surveillance of relatively small populations.

Perez Gutthann and Garcia Rodriguez conducted a nested case-control study of the risk of hospitalization for acute liver injury among users of NSAIDs who were concomitantly taking other drugs known to be associated with hepatic injury [14]. The source population consisted of more than 200,000 residents of the province of Saskatchewan who had filled one or more prescriptions for one of five commonly prescribed NSAIDs during the period 1982-86. Crude estimates of additional risk from use of other medications ranged from $1/10^5$ prescriptions for methyldopa or ampicillin to $12/10^5$ prescriptions for erythromycin estolate. No hospitalizations occurred with more than 70,000 prescriptions for other salts of erythromycin. The authors concluded that concomitant use of two or more drugs

associated with hepatic injury increased the risk above that which would be calculated from the sum of the individual risks. Among those not currently taking any drugs they estimated a baseline rate of $3/10^5$ person-years. This was increased by approximately 2.1 fold among those taking NSAIDs and no other known hepatotoxic drugs and by approximately 11.6 fold among those taking both NSAIDs and one or more other prescription medications associated with liver injury from the literature.

Action Item: Additional research in defining the background rate of specific hepatic syndromes (including hepatic failure) in specific populations of users is required. Population-based studies may be helpful in defining 1) the incidence of drug associated and non-drug associated hepatotoxicity, 2) the comparative incidence and rates of drug associated injury, 3) the natural history/clinical course and clinical spectrum of hepatotoxicity, 4) the morbidity and mortality associated with drug induced hepatotoxicity, and 5) the possible risk factors for drug associated hepatotoxicity. This is offered as a topic for discussion at the workshop.

Initiative 4: Quantifying the Frequency of LFT Abnormalities and Evaluating the Usefulness of Liver Monitoring Recommendations in Preventing Drug Injury.

Recently Duh et al quantified the incidence of liver function abnormalities in the general population of Massachusetts in a study using computerized data files from a health maintenance organization [15]. Liver enzyme abnormalities were defined by a set of international consensus criteria based on serum liver enzymes[5]. Among several etiologic categories of liver enzyme abnormalities (drug-associated, biliary pathologies, mononucleosis, alcohol, malignancy, viral hepatitis (A, B, C), or unknown), drug-associated abnormalities were most common. The incidence was 40.6 persons per 100,000 persons per year, with a 95% confidence interval of 29.3 – 51.8, based on a total of 50 cases. This study evaluated only outpatient use of prescription drugs, including non-steroidal anti-inflammatory drugs (NSAIDs), lipid-lowering agents, isoniazid, methotrexate, oral erythromycin, sulfa drugs, and chemotherapeutic drugs.

The best source of information on the frequency of liver function abnormalities among users of specific drugs comes from long-term controlled clinical trials. For example, in the Scandinavian Simvastatin Survival Study (4S), 4,444 patients with a history of angina pectoris or acute myocardial infarction but without baseline impairment in hepatic function were randomized to simvastatin 20-40 mg daily (N = 2,221) or placebo (N = 2,223) for a median follow-up of 5.4 years[16,17]. During the trial ALT and AST were measured every six weeks during the first eighteen weeks and every six months thereafter. Twenty-four patients in the placebo group (1.1%) and 20 in the simvastatin group (1.0%) had levels of AST > 3 times the upper limit of normal (ULN) at least once, while 7 patients in the placebo group (0.3%) and 5 in the simvastatin group (0.2%) had levels of AST > 3X ULN on more than one occasion. Thirty-two patients in the placebo group (1.6%) and 46 in the simvastatin group (2.2%) had levels of ALT > 3 X ULN at least once, while 12 patients in the placebo group (0.6%) and 14 in the simvastatin group (0.7%) had levels of ALT > 3 X ULN on more than one occurrence. The protocol called for discontinuation of patients with increases in AST or ALT above 4 times the ULN confirmed by a repeated measurement. Five patients in the placebo group (0.2%) and 8 in the simvastatin group (0.4%) were discontinued because of elevated liver enzymes. There were 2 cases of non-viral hepatitis in the placebo group and 1 case in the simvastatin group. These results provide information on the frequency of aminotransferase elevations in patients without baseline impairment of hepatic function treated with simvastatin or placebo for more than five years. Overall it remains difficult to conclude that simvastatin caused any injury. It is clear, however, that a substantial rate of low-level transaminase increases can be expected in untreated patients. There was no evidence of drug-induced hepatitis.

Analyses of the placebo groups of controlled clinical trial databases can be useful in understanding the background incidence of sustained aminotransferase elevations (disease-specific intrasubject variability) not associated with drug-induced hepatic dysfunction. This information would be valuable in setting the ‘cut points’ for comparing clinical laboratory values in treated versus placebo patients and critical for evaluating LFT

abnormalities in the long-term, uncontrolled, extension periods in clinical trials. There is a dearth of published information on this subject (i.e. intrapatient variability over time in liver function tests) and most of what is published examines this variability over short periods (e.g. 10 days to 2 weeks) in healthy populations [18]. Some industrial pharmaceutical researchers have used these types of data to establish 'reference ranges' for the analysis of clinical laboratory data. Interestingly, these data from controlled trials reveal sharp contrasts between the derived 'reference ranges' and the usual 'health-associated' reference values with the derived reference ranges considerably wider. Demographic and other covariates (tobacco and ethanol use) were examined but there was no attempt to derive disease-specific reference ranges, a potential next step [19].

Whether periodic liver monitoring is useful in preventing serious drug-induced hepatic injury is not established. The US Preventative Services Task Force has outlined two criteria for judging the appropriateness of using a test to screen for disease [20].

- The test must be able to detect the target condition (in this case, serious liver injury) earlier than without screening and with sufficient accuracy to avoid producing large numbers of false-positive and false-negative results (i.e. *accuracy*).
- Screening for and treating persons with early disease should improve the likelihood of favorable health outcomes (e.g. reduced disease-specific morbidity or mortality) compared to treating patients when they present with signs or symptoms of the disease (i.e. *effectiveness of early detection*).

Test accuracy depends not only on sensitivity and specificity, but also on how common the disease is in the population screen. Even when a test is highly sensitive and specific, nearly all abnormal tests will be false positives if the disease is rare. For example, even if LFTs were 100% sensitive and 99% specific in identifying serious hepatic injury, about 98% of all positive tests would be false-positives if serious hepatic injury occurred in only one in every 5,000 patients treated (see Figure 1). Thus, when periodic liver monitoring is recommended for a drug where hepatic injury is infrequent, not only will thousands of tests be required to avert one true case of hepatic injury, but also nearly all positives will be false positives. Both factors combine to make periodic liver monitoring highly inaccurate, very expensive and burdensome to patients, physicians and the healthcare system. Patients may be needlessly taken off effective therapy and subjected to unnecessary evaluations.

Apart from the issue of predictive value, the effectiveness of a liver monitoring program depends on the extent to which early detection of liver test abnormalities and prompt drug withdrawal can prevent development of serious hepatic damage. Before recommending periodic liver monitoring in drug labeling, there should be reasonable assurance that it is likely to reduce the risk of serious hepatic injury and the amount of testing required to avert one case needs to be estimated. With this information, one could then determine whether this screening would represent an appropriate use of medical resources either in economic terms (i.e. cost-effectiveness) or simply in terms of resources used per case averted [21].

Action Item: Disease-specific, long-term, intrasubject and intersubject variability in transaminase levels should be assessed from the placebo arms of long-term controlled trials in various populations. In addition, further research on the role and value (from an individual and population perspective) is required. These topics are offered as a potential topic for discussion at the workshop.

Initiative 5: Active and Sentinel Surveillance Strategies

FDA is exploring the application of active surveillance techniques to enhance reporting and determine the incidence rates of drug related adverse events.
Of particular relevance to the drug-associated liver injury are databases that might be used to identify all, or a large fraction of, serious liver injuries. Patient histories could then be reviewed for possible drug exposure. Possible databases include data of the United Network of Organ Sharing (UNOS) funded by the HRSA, death certificates reporting liver related deaths derived from NCHS database, and the National Electronic Injury Surveillance System (NEISS) sponsored by the Consumer Product Safety Commission (CPSC). UNOS is a data collection system for national transplantation registration and organ distribution. One possible approach would be to add a component to more fully capture drug exposure history for liver transplant registrants. Potentially, the group of patients with the highest probability of providing evidence of drug-induced etiologies for their hepatic injury may be the approximately 200 cases/year of idiopathic liver failure. An advantage of this approach is that the data are collected in real-time upon registration of the patient. NEISS is funded by the Consumer Product Safety Commission and has proven successful in identifying acute injury presenting to emergency rooms due to consumer products. It collects data from a nationally representative sample of 101 hospital-based emergency rooms. It may be possible to expand the effort to allow determination of drug exposure histories for designated diagnoses of interest such as acute liver failure. From this, a population-based incidence rate for acute liver failure due to specific drugs might be ascertained from this representative sample. The Acute Liver Failure Study Group network, if expanded, may also represent an opportunity to leverage an existing system for surveillance purposes [6].

Action Item: The possibility of establishing and funding a sentinel system for detecting drug-induced hepatic injury should be discussed at the workshop. Such as system should be designed to capture as much of the universe of severe hepatic injury (e.g. patients going to transplant) and be able to obtain detailed medical and medication histories as well as injured liver specimens for research on the mechanism of drug-induced hepatotoxicity. At the very least, a retrospective analysis of drugs who are known to rarely cause severe hepatotoxicity (e.g. troglitazone, bromphenac, trovofloxacin, etc.) should compare the efficiency (cost and rapidity) and informativeness (completeness and detail) of the spontaneous reporting system and a system like UNOS or the ALF network. Before investing in what may be an labor and financially intensive effort such as these sentinel systems, it would be important to see if they find things sooner with more and better information upon which to base decision making.

Initiative 6: Using Post-approval Experience to Improve the Predictive Value of Preclinical and Clinical Experience (cataloging of FDA and industry liver experience)

Although it has not been possible to do so fully to date, pre-clinical and clinical study data would, ideally, identify significant hepatotoxins prior to marketing. By systematically comparing post-approval information on drug-related liver disease with information available from preclinical studies and pre-approval clinical trials, it may be possible improve the ability of these data to serve as a reliable predictor of hepatotoxicity. This potential is considered in detail in the Clinical White Paper. To make such an exercise meaningful one would need to examine a sizeable sample of drugs that did and did not have post-approval evidence of important hepatic toxicity.

Action Item: Such a project is clearly feasible and should be discussed at the workshop (see Clinical White Paper for additional suggestions related to this topic).

Initiative 7: Nationwide Research Proposal to Determine Mechanisms of Serious Drug-induced Hepatic Injury.

Most serious hepatotoxins have shown evidence of mild hepatocellular injury observed in clinical trials (manifested by ALT elevations in a proportion of recipients and sometimes by elevated bilirubin). It seems likely that mild injury is mechanistically related to life-threatening liver injury seen after marketing, but it is not clear what factors cause injury to be much more serious in a small fraction of patients. Although the molecular events that result in elevations in serum ALT may be necessary for severe injury, they are apparently not sufficient to produce life-threatening liver injury because the risk of life-threatening liver injury is generally several orders of magnitude lower than that for ALT elevations (based on clinical trial data). For example, troglitazone produced ALT elevations >3 X ULN in 1 in 50 patients in clinical trials, but carried a lower risk of life-threatening liver injury patients after marketing. Furthermore, some drugs such as taccrine are well known to cause frequent and large elevations in aminotransferase levels without significantly increasing the risk of acute liver failure. The reason for this phenomenon is not well understood and the challenge of differentiating *bona fide* hepatotoxins from those causing 'benign' aminotransferase elevations prior to marketing is a major problem in pre-marketing assessment (see Clinical White Paper).

It seems that it may prove possible to discover the molecular/genetic basis for the different responses to hepatotoxins: no effect at all, mild hepatotoxicity, or serious injury. Pharmaceutical companies have a great deal of data from clinical trials on mild injury and genetic material from patients with mild injury could be saved and studied in order to determine patterns of individual susceptibility. Some companies are already doing this. The more difficult material to obtain will be from patients outside of trials who have suffered severe hepatotoxicity reactions to a drug. These patients are particularly important because they, presumably, have the specific characteristics associated with severe injury.

A potential solution to this difficulty might be to create a nation-wide severe drug induced liver disease data and tissue bank accessible to investigators for hypothesis testing. This might be a freestanding initiative or, alternatively, it could be added on to an existing effort (e.g. the Acute Liver Failure Study Group).

As it is important to discover hepatotoxicity to a wide range of drugs, the database and tissue bank probably should be under the auspices of the FDA or some other organization such as the NIH or AASLD. Access to the tissue and data will need to be tightly controlled by an advisory committee whose members have minimal conflict of interest.

Ideally, information on patients would be obtained as close to the time of injury as possible, probably before the case is even reported to a manufacturer or FDA. A standardized comprehensive information data sheet should be prepared for each subject, including a sufficiently detailed medical history, a record of all liver function tests performed (values and dates), and concomitant medications in the time period leading up

to the event. Whole blood should also be obtained from patients; privacy concerns will need to be addressed. Lymphocytes will be isolated, immortalized, and frozen. This will provide an inexhaustible supply of genomic DNA, and will also provide functioning immune cells for function and protein analysis. Plasma should also be saved, primarily for searches for anti-liver antibodies, but analysis of other proteins might be helpful.

It should be noted that liver tissue would not be required or obtained for the aforementioned tissue bank because:

- the small size of percutaneous or transjugular biopsy material generally means that the entire tissue specimen would be required for diagnostic purposes (i.e., the reason for the biopsy)
- liver biopsies can generally not be obtained solely for research purposes because of the high morbidity associated with the procedure and
- an endstage liver (as would be expected from those patients awaiting transplant) is unlikely to be useful for research purposes.

Action Item: The scientific and practical feasibility and ethical/privacy issues surrounding this proposal is offered as a discussion topic for the workshop.

Initiative 8: Hepatotoxicity Management Procedures Initiative

Once significant, unanticipated drug-induced hepatotoxicity has been identified and its frequency quantified as well as possible, there needs to be a consistent approach to managing and communicating the risk. The foundation must be carefully laid for such an effort. FDA and the pharmaceutical industry could analyze the consistency of labeling representation of hepatotoxicity identified after initial marketing. This should include detailed analysis of the location, prominence, and content of the labeling information and a retrospective analysis of the data that were relied upon for labeling.

There should also be an effort to quantify the effects of risk-management strategies on prescribing behavior. Awareness of and compliance with labeling changes and/or 'Dear Dr' letters can be measured using survey methodology and population-based estimates of drug use. For example, a study quantifying the frequency of dispensing of contraindicated drugs with terfenadine that quantified the impact of serial risk communication efforts and revealed residual, recalcitrant coprescribing was important in the overall risk-benefit assessment of that drug [22]. Similarly, assessment of compliance with liver monitoring recommendations should be considered for all drugs whose safe use is contingent upon on serial monitoring of liver chemistries.

Action Item: A concerted research effort should undertake the evaluation of different risk communication/risk management strategies. The overall objective of such an effort would be to try to understand what works and what does not in attempting to modify prescribing and consumer behavior. This proposal is offered as a discussion topic for the workshop.

Summary: Eight action items are proposed for further discussion at the PhRMA/FDA/AASLD Workshop.

These are:

1. Development of a Standardized Nomenclature and Definitions for Liver Test Abnormalities and Liver Injury as well as a Standardized Approach to Causality Assessment in Postmarketing Reports.
2. Passive surveillance: improvements increased and improved reporting, and data analysis techniques.
3. Improved Ascertainment of Background Rates of Elevated Aminotransferases and Hepatic Injury
4. Quantifying the Frequency of LFT Abnormalities and Evaluating the Usefulness of Liver Monitoring Recommendations in Preventing Drug Injury.
5. Active and Sentinel Surveillance Strategies
6. Using Post-approval Experience to Improve the Predictive Value of Preclinical and Clinical Experience (cataloging of FDA and industry liver experience)
7. Nationwide Research Proposal to Determine Mechanisms of Individual Susceptibility to Serious Drug-induced Hepatic Injury.
8. Hepatotoxicity Management Procedures Initiative

Table 1. Definitions and Types of Liver Inquiry (Strawman)

Liver Injury	Hepatocellular	Cholestatic	Mixed
>2-3xULN of ALT (SGPT) OR	>2-3 x ULN in ALT and nl Alk Phos OR	>2 x ULN in Alkaline Phosphate OR	>2-3 x ULN ALT AND >2 x ULN Alkaline Phosphate OR
>2 x ULN conjugated Bilirubin, OR elevated AST (SGOT), Alkaline Phosphate and Total Bilirubin (one of these must be >2 x ULN)	Ratio of ALT to Alkaline Phosphate ≥ 5	Ratio of ALT to Alkaline Phosphate ≤ 2	Ratio of ALT TO Alkaline Phosphate AND Ratio of ALT and Alkaline Phosphate between 2 and 5

Table 2 Sample Proposed Formats for Collecting Hepatotoxicity Data

LIVER INJURY

Signs of severe injury include a marked elevation of ALT or conjugated bilirubin (CB), PT prolongation, the presence of jaundice in association with hepatocellular injury, or the presence of hepatic encephalopathy.

1 Abbreviated List

Shaded cells: most desired data elements.

1. RELEVANT HISTORY AND CLINICAL CONDITIIONS					
Hepatobiliary disorder	No Yes Specify	Alcohol abuse	No Yes Specify	Risk factor(s) for viral hepatitis	No Yes Specify

2. RELEVANT TESTS/LABORATORY DATA (where ND=not determined)							
a). Lab. Tests	Normal Range	**Before** treatment	**During treatment**		**After cessation of the suspected drug** (enter only lowest levels observed) Drug(s) discontinued: _____		
			Earliest Abnormalities Observed*	Highest level observed	Day 0	First 8 days	Day 9-180
DATE (M/D/Y)							
ALT (SGPT)							
AST (SGOT)							
AP (Alk Phos)							
T. Bili. (TB)							
C. Bili. (CB)							
Protime (PT)							
GGT							
CPK							
b).Serology	**Not done**	**Absent**	**Present**	**Titer**	**Date M/D/Y)**	**c) Liver Biopsy**	
Anti HAV/IgM Anti HBc/gM Anti-HCV Other serology						Not done Done Findings	

Table 2 *(Continued)*

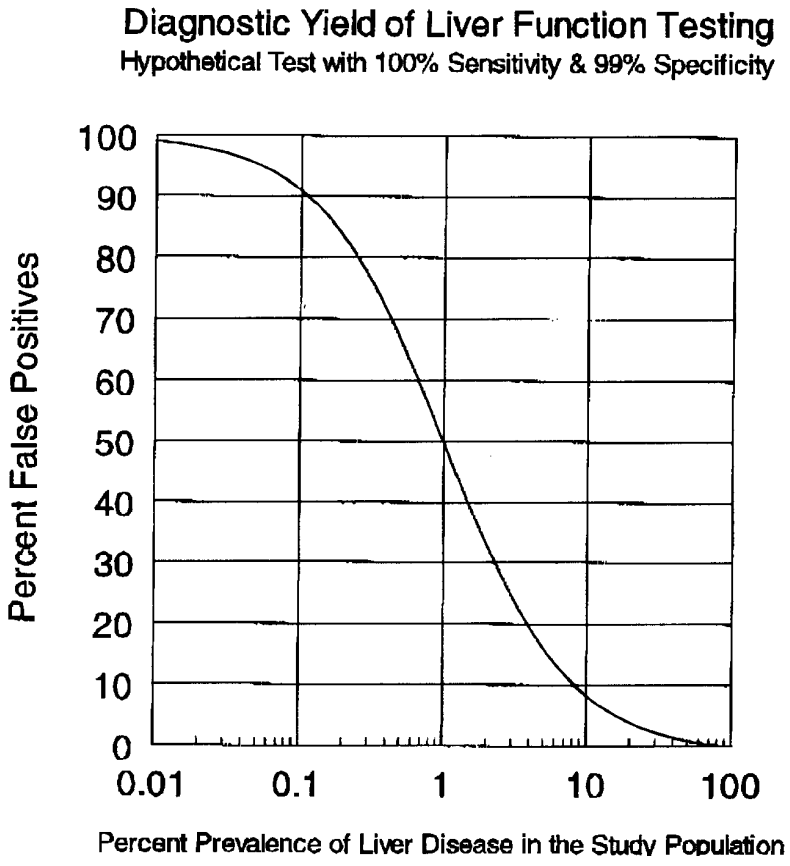
2 Comprehensive List

1. RELEVANT HISTORY AND CLINICAL CONDITIIONS						
	No Yes Specify		No Yes Specify		No Yes Specify	
Hepatobiliary disorder		Right side heart failure		Occupational toxic agent		
Alcohol abuse		Recent hypotension		Intravenous drugs abuse		
Drug allergy		Cancer		Acupuncture		
Auto-immune		Transfusion of blood products		Recent travel to Africa, Asia		

2. RELEVANT TESTS/LABORATORY DATA (where ND=not determined)							
a). Lab. Tests	Normal Range	**Before** treatment	**During treatment**		**After cessation of the suspected drug** (enter only lowest levels observed) *[re lowest – why- makes no sense]* Drug(s) discontinued: _____		
			Earliest abnormalities observed*	Highest level observed	Day 0	First 8 days	Day 9-180
DATE (M/D/Y)							
ALT (SGPT)							
AST (SGOT)							
AP (Alk Phos)							
T. Bili. (TB)							
C. Bili. (CB)							
Protime (PT)							
GGT							
CPK							
b).Serology	**Not done**	**Absent**	**Present**	**Titer**	**Date M/D/Y)**	**c) Liver Biopsy**	
Anti HAV/IgM Anti HBc/IgM Anti-HCV Anti-CMV Igm Anti-nuclear Anti-native DNA Anti-smooth muscle Anti-mitochondria Other serology						Not done Done Findings_____ *[Re not done, done findings – what's this mean?]*	

*Enter only result from the same day

Figure 1:


Diagnostic Yield of Liver Function Testing
Hypothetical Test with 100% Sensitivity & 99% Specificity
Percent False Positives
100
90
80
70
60
50
40
30
20
10
0
0.01
0.1
1
10
100
Percent Prevalence of Liver Disease in the Study Population

References:

1. International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). The Extent of Population Exposure to Assess Clinical Safety. ICH-E1A. Available at: http://www.ifpma.org/ich1.html.
2. Tilson H. Pharmacoepidemiology in the Pharmaceutical Industry. In: Detection of New Adverse Drug Reactions. 1998.
3. Benefit - Risk Balance for Marketed Drugs: Evaluating Safety CIOMS IV, Geneva 1998
4. Danan G. Liver Test Abnormalities. In Adverse Drug Reactions. A Practical Guide to Diagnosis and Treatment, Benichou (Ed), John Wiley and Sons Ltd. 1994.
5. Benichou C, et al. Criteria of drug-induced liver disorders: Report of an international consensus meeting. Journal of Hepatology, 1990:11:272-276.
6. Lee WM. Acute Liver Failure. New England Journal of Medicine 1993; 329(25): 1862-1872.
7. Hoofnagle JH et al. Fulminant Hepatic Failure: Summary of a Workshop. Hepatology 1995; 21:240-252.
8. DuMouchel W. Bayesian Data Mining in Large Frequency Tables, With an Application to the FDA Spontaneous Reporting System. The American Statistician.53:177-190,1999
9. Guess HA. How should acute hepatic drug effects be studied epidemiologically? Epidemiology. 1993; 4:487-489.
10. Walker AM et al. The infrequency of liver function testing in patients using nonsteroidal anti-inflammatory drugs. Archives of Family Medicine 1995; 4:24-29.
11. Beard K et al. Outpatient drug-induced parenchymal liver disease requiring hospitalization. Journal of Clinical Pharmacology. 1986; 26:633-637.
12. Walker AM, Cavanaugh RJ: The occurrence of new hepatic disorders in a defined population. Post Marketing Surveillance. 1992; 1:107-117.
13. Garcia Rodriguez LA et al. The role of non-steroidal anti-inflammatory drugs in acute liver injury. British Medical Journal 1992; 305:865-868.
14. Perez Gutthann S and Garcia Rodriguez LA: The increased risk of hospitalizations for acute liver injury in a population with exposure to multiple drugs. Epidemiology 1993; 4:496-501.
15. Duh MS et al. Descriptive epidemiology of acute liver enzyme abnormalities in the general population of central Massachusetts. Pharmacoepidemiology and Drug Safety. 1999; 8:275-283.
16. Scandinavian Simvastatin Survival Study Group: Randomized trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). Lancet. 1994; 344(8934): 1383-9.
17. Pedersen TR, et al. Safety and tolerability of cholesterol lowering with simvastatin during 5 years in the Scandinavian Simvastatin Survival Study. Archives of Internal Medicine 1996; 156(18): 2085-92.
18. Rosenzweig P et al. Transaminase elevations on placebo during phase I trials: prevalence and significance. British Journal of Clinical Pharmacology 1999; 48:19-23.

19. Wilson MG et al. Lilly reference ranges. In: Encyclopedia of Biopharmaceutical Statistics In press.
20. Preventative Services Task Force. In: Guide to Clinical Preventative Services. Second Edition. Williams & Wilkins. Baltimore 1996.
21. Guess HA: How should acute hepatic drug effects be studied epidemiologically? Epidemiology 1993; 4:487-489.
22. Burkhart G et al. Temporal decline in filling prescriptions for terfenadine closely in time with those for either ketoconazole or erythromycin. Clinical Pharmacology and Therapeutics 1997; 61:93-6.]